

**ECON 211B**  
**Winter 2017**

**Econometrics**  
**Homework 5**

Please estimate the effects of the MLDA on alcohol consumption and on mortality. The dataset **NHIS.csv** is a sample drawn from the National Health Interview Sample Adult Files 1997-2007. It includes the following variables: *HS\_diploma*, *Hispanic*, *white*, *black*, *uninsured*, *employed*, *married*, *workin\_lw*, *going\_school*, *male*, *days\_21* (Days to 21st birthday), *drinks\_alcohol* (reports they drink alcohol), *AGE\_yrs*, *perc\_days\_drink* (percent of days on which they report drinking). Each row on the dataset is a record for an individual. The file **mortality data.csv** contains mortality rates by age overall and broken down by primary cause of death.

1. Create a carefully labeled figure of the first stage relationship between the MLDA and whether or not people drink alcohol (variable name: *drinks\_alcohol*). You will need to make age cells. Experiment with different bin sizes what size did you end up with and why?
2. Create a table of regression estimates of the increase in the proportion of people who drink (variable name: *drinks\_alcohol*) with different order polynomials going from least to most flexible. Use the *rdrobust* package to get a starting point for the optimal bandwidth which is available at <https://sites.google.com/a/umich.edu/rdrobust/implementation-in-stata>. Assess the quality of the fit by superimposing the fitted values of the regression over the scatter plot from question 1.
3. Check if the regression estimates are robust to bandwidth choice by estimating them for 20-30 bandwidths around your chosen one and plotting them with bandwidth on the x axis and estimate and confidence interval on the right.
4. Check if there is any evidence of sharp changes in characteristics of the sample at 21. Make a table that cleanly presents these estimates.

5. Do a crude density check by examining and checking for a change in the number of people responding to the survey at age 21.
6. Does the MLDA reduce drinking? If so by how much.
7. Make a figure with the age profile of deaths due to all causes.
8. Make a figure with deaths due to Motor Vehicle Accidents (MVAs).
9. Are these estimates robust to bandwidth choice? Document this?
10. Does the MLDA reduce death rates?
11. Estimate the effect of drinking on MVA. Use a two sample instrumental variables approach (reduced form/first stage). Estimate the standard errors using the delta method.
12. Do you believe the assumptions under which the MLDA is a valid instrument for alcohol consumption are met. Assess them.